

Nina Berend/Rudolf Schmidt

Perspektiven der computergestützten Untersuchung von Variation im gesprochenen Deutsch

1. Das IDS-Forschungsprojekt: Eckdaten

Im Institut für Deutsche Sprache in Mannheim wird zur Zeit ein neues Forschungsprojekt entwickelt, das sich mit Untersuchung der Variation im gesprochenen Deutsch beschäftigt.¹ Da es sich um die Bearbeitung größerer Korpora handelt, wird dieses Projekt computergestützt und mit Einsatz von Verfahren der automatischen Sprachanalyse durchgeführt werden. In den letzten Jahren wurden im Institut große Bemühungen für Computereinsatz in der Forschung unternommen und es hat sich dort bereits eine korpus technologische Tradition insbesondere bei der gesprächsanalytischen Forschung gebildet (vgl. Kallmeyer 1997). Das hier vorgestellte Projekt *Aussprachevariation* soll diese Tradition fortsetzen und um die phonetische Dimension erweitern. Das Ziel des Projekts ist die Analyse und Dokumentation der Aussprachevariation in den verschiedenen Sprachvarietäten des Deutschen, vor allem in der deutschen Standardsprache und in den regionalen Umgangssprachen. Die empirische Basis bilden die umfangreichen IDS-Korpora des gesprochenen Deutsch²; die korpus technologische Basis ist das Alignment (Text-Ton-Synchronisation) und COSMAS II³ (vgl. dazu ausführlich Abschnitt 4). Angestrebtes Ergebnis ist eine mit COSMAS II verknüpfte, korpusbasierte Datenbank von Wörtern, mit Zugriff auf Kontexte, Angaben zu Kontext- und Situationstypen, Charakterisierungen von Standard- oder Regiolekteigenschaften sowie von Formalität bzw. Informalität des Sprechens. Die zentrale Fragestellung des Projekts bezieht sich auf die Aussprache des Deutschen als soziolinguistisches Phänomen. Es geht um die Varianten der Aussprache und die Verteilung der variablen Aussprachemerkmale auf verschiedene Regionen und verschiedene Textsorten bzw. Kommu-

¹ Das Projekt ist in der Abteilung Pragmatik des Instituts angesiedelt. Projektgruppe (bisher): Karl-Heinz Bausch, Nina Berend, Werner Kallmeyer (Leitung), Rudolf Schmidt (vgl. Bibliographie: Projektbeschreibung, 2001).

² [<http://www.ids-mannheim.de/kt/corpora-ges.html>].

³ [<http://www.ids-mannheim.de/zdv/cosmas2/>].

nikationssituationen. Die verschiedenen Ausprägungen der Aussprachevarianten des Deutschen (Eisenberg 1998) sollen empirisch untersucht und linguistisch und soziolinguistisch beschrieben werden.

Abb. 1 stellt schematisch das Rahmendesign des Projekts dar. Sehr verallgemeinert besteht das Projekt aus drei Komponenten. Die links dargestellte Komponente bildet die empirische Basis und Materialgrundlage des Projekts. Es handelt sich um die im Institut bereits vorliegenden Aufnahmen des gesprochenen Deutsch. Der mittlere Bereich des Schemas stellt die zweite Grundkomponente des Projekts dar: die korpus-technologische Basis, die eine Grundvoraussetzung zur Durchführung dieses Projekts ist (Verfahren, mit denen die Korpora in Bezug auf Varianten automatisch analysiert werden). Die dritte Komponente stellt das zentrale angestrebte Ergebnis des Projekts dar – die Datenbank von Aussprachevarianten des Deutschen. Die einzelnen Komponenten des Projekts werden im Folgenden kurz skizziert (vgl. Abschnitte 2, 3 und 4).

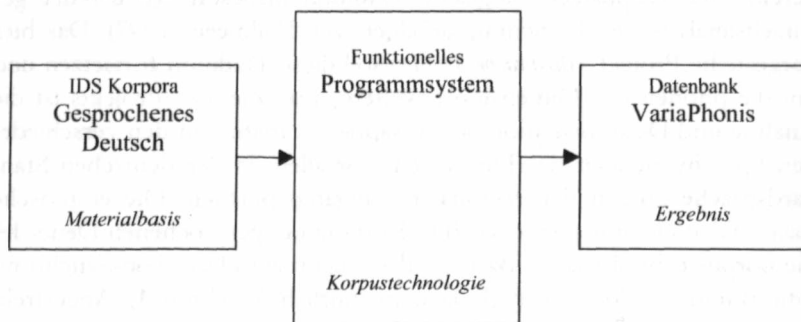


Abb. 1: Allgemeines Rahmendesign des Projekts

2. Empirische Basis des Datenbank-Projekts

Die Ausgangsbedingung für die Realisierung des geplanten wissenschaftlichen Vorhabens sind die im Institut vorliegenden Korpora des gesprochenen Deutsch, die den Sprachgebrauch von mehreren Tausend Sprechern des Deutschen in verschiedenen kommunikativen Situationen repräsentieren. Aus verschiedenen Gründen sind die Aufnahmen gut

geeignet, als empirische Basis für die geplante Datenbank verwendet zu werden. Es seien hier nur einige davon kurz erläutert.

Umfangreiches Korpus: Die IDS-Korpora stellen gegenwärtig die größte Sammlung der Aufnahmen des gesprochenen Deutsch im gesamten deutschsprachigen Gebiet dar. Es liegen bereits 32 Korpora mit insgesamt über 15.000 Aufnahmen vor (vgl. Anm. 2). Die Integration von einigen weiteren Korpora steht unmittelbar bevor. Die vorliegenden Aufnahmen bilden eine ausreichende Basis für die Untersuchung der Aussprachevariation.⁴

Breites Textsortenspektrum: Bei den IDS-Korpora handelt es sich um Sprachaufnahmen in verschiedenen Kommunikationssituationen bzw. Kontextstilen oder Textsorten. Das Textsortenspektrum umfasst z.B. Beratungen, Fragen, Gedichte, Gespräche, Gremiensitzungen, Inhaltsangaben, Interviews, Kommentare, Lesungen, Märchen, Nachrichten, Predigten, Pressekonferenzen, Reportagen, Schlichtungsgespräche, Statements, Talkshows, Unterhaltungen, Vorträge u.a. Für die Analyse der aktuellen Variationsverhältnisse des Deutschen ist es notwendig, ein breites Spektrum an Textsorten als empirische Basis mit einzubeziehen. Der Atlas der Aussprache des Schriftdeutschen, die zur Zeit sicherste Quelle für die Aussprachevariation in der Bundesrepublik (König 1989), beruht bekanntlich lediglich auf Daten der Vorlesesprache.

Tatsächlicher Sprachgebrauch: Die Korpora sind repräsentativ für das gesprochene Deutsch in authentischen Kommunikationssituationen, da es sich zum großen Teil um Aufnahmen handelt, die nicht speziell zu linguistischen Zwecken erstellt worden sind, wie z.B. Beratungen, Talkshows, Schlichtungsgespräche u.a. Sie dokumentieren den realistischen Sprachgebrauch in realen, tatsächlich stattgefundenen Kommunikationssituationen. Aufnahmen für linguistische Zwecke sind jedoch solche, die in extra für diese Zwecke organisierten Kommunikationssituationen entstehen.

Aufbereitungsstufe: Die dem Projekt zur Verfügung stehenden Aufnahmen sind bereits zum großen Teil aufbereitet und digitalisiert und somit für die automatische Analyse ohne beträchtlichen Aufwand verwendbar, was aus ökonomischen Gründen für die Durchführbarkeit des Projekts nicht irrelevant ist.

⁴ Neue Datenerhebungen sind nur im Rahmen von gezielten Zusatzaufnahmen für bestimmte Projektzwecke vorgesehen.

3. Aussprachedatenbank VariaPhonis

Die im Schema (Abb. 1) rechts angeordnete Komponente stellt das zentrale angestrebte Ergebnis des Projekts dar: die Datenbank *VariaPhonis*. Die Datenbank soll das leisten, was die vorliegenden Aussprachewörterbücher nicht tun: Die Vermittlung von systematischen Informationen über die Aussprachevarianten und Aussprachevarietäten des Deutschen. Als Informations- und Recherchesystem ist sie vor allem in der Auslandsgermanistik und im Unterricht Deutsch als Fremdsprache einsetzbar, aber auch für die Optimierung von Wörterbüchern der Aussprache des Deutschen und für die Analyse aktueller Variation und ihrer Zusammenhänge mit Sprachwandel. Als Informations- und Rechercheinstrument wird die Datenbank im Endausbau *online* zugänglich sein. Es ist vorgesehen, dass die Datenbankstruktur folgenden Aufbaukriterien entspricht:

- korpusbasiert,
- kontextorientiert,
- mit expliziten Informationen zum Variantengebrauch,
- mit vernetzter Struktur der Informationseinheiten.

Betrachten wir die einzelnen Aufbaukriterien etwas näher. Bei der Korpusbasiertheit handelt es sich darum, dass in der Datenbank grundsätzlich nur solche Varianten dokumentiert werden, die im Korpus auch tatsächlich belegt sind. Dadurch wird angestrebt, dass die Datenbank Auskunft gibt über den realistischen Sprachgebrauch und Gebrauch von Aussprachevarianten, die im gesprochenen Deutsch auch tatsächlich vorkommen (z.B. auch über Varianten, die die regionalen Standards des Deutschen charakterisieren und die in Aussprachewörterbücher nicht aufgenommen werden).⁵

Das nächste Aufbaukriterium, die Kontextorientiertheit, gewährleistet den Zugriff auf die Kontexte aller belegten Varianten. Es handelt sich dabei sowohl um den schriftlichen Kontext der Textdatei als auch um das unmittelbar dazugehörige Sprachsignal. Die Varianten werden somit nicht isoliert dargestellt, wie das in Wörterbüchern der Fall ist (wie z.B. bei der Darstellung der Umgangs lautung, vgl. Duden 2000), sondern in ihren Gebrauchskontexten. Dadurch wird den Nutzern die Möglichkeit gegeben, die Verwendung der Aussprachevarianten unmittelbar in den

⁵ Vgl. z.B. Duden 2000, Vorwort: „[...] Varianten bleiben ausgeblendet oder auf ein Mindestmaß beschränkt“.

dokumentierten Kontexten zu recherchieren. Das eröffnet für Sprachwissenschaftler, insbesondere auf dem Gebiet der Variationslinguistik, eine Perspektive zur Erforschung der Gebrauchskontexte von Varianten, wie sie bisher nicht vorhanden war.

Das dritte Aufbaukriterium stellt einen weiteren wesentlichen Unterschied zu den bereits vorliegenden Aussprachewörterbüchern dar. Die dokumentierten Aussprachevarianten werden durch Angaben ergänzt, die die spezifischen Gebrauchsaspekte der Variante verdeutlichen. Die einzelnen Komponenten der Angaben nennen wir *Deskriptoren*. Durch ein Deskriptorensystem werden explizite Erklärungen zum Variantengebrauch bereitgestellt, die im Endausbau der Datenbank das soziolinguistische Informationssystem über Aussprachevarietäten darstellen. Dieses letzte Kriterium macht deutlich, dass im Rahmen des Projekts versucht wird, neue Beschreibungsformate für die Varianten des gesprochenen Deutsch zu entwickeln. Die Abfragemöglichkeiten werden sich nicht beschränken auf die Angaben über die Ausspracheformen einzelner Laute oder Wörter, wie bisher in Wörterbüchern üblich, sondern darüber hinaus wird es möglich sein, auch andere relevante Informationen abzufragen, wie z.B. Angaben über Gebrauchskontexte der Varianten, über Sprachlagen, Varietäten, Eigenschaften von Sprachlagen und Varietäten, usw.

Die vernetzte Struktur der Informationseinheiten schließlich, das letzte der oben genannten Aufbaukriterien, soll flexible Möglichkeiten der Informationsauswahl und Informationstiefe bieten. Jeder Nutzer kann durch diese vernetzte Struktur (und das entsprechende Linksystem) die gewünschten Informationskriterien individuell bestimmen.

4. Automatische Detektion und Verwaltung von Varianten

Die mittlere Komponente des Projekts (vgl. Abb. 1) ist die Korpustechnologie bzw. die technologische Grundlage, die die automatische Analyse der Korpora ermöglicht. Aufgrund des umfangreichen Materials ist es – wie oben schon erwähnt – nötig, das Aufspüren von Sprachvarianten maschinell durchzuführen. Dies wäre eine Aufgabe für einen um eine Vielfalt von Sprachvarianten angereicherten Spracherkenner. Nach dem bisherigen Stand der Technik ist jedoch bei der Qualität der Sprachaufnahmen nicht damit zu rechnen, dass ein Spracherkenner Ergebnisse in ausreichender Qualität liefern kann. Da jedoch ein Großteil der IDS-Korpora verschriftlicht ist, kann ein solches Programm auf solche In-

formationen zurückgreifen, was erwarten lässt, dass sich die Qualität der Ergebnisse verbessert. Ein derartiges Programm ist am IDS im Rahmen eines Projekts mit der Bundeswehr entwickelt worden. Es heißt SPRAT (SPeech Recognition and Alignment Tool) und gestattet sowohl die Synchronisation von Text und Ton (sogenanntes Alignment) als auch Spracherkennung (vgl. Schmidt 2001). Wie ein solches Programm aufgebaut ist und durch welche Modifikationen es in die Lage versetzt werden kann, auch Varianten aufzuspüren, ist Gegenstand der folgenden beiden Unterkapitel.

4.1. Technologische Grundlagen

Die Aufgabe eines Aussprachevariantendetektors nach oben gemachter Erläuterung ist es – bei Vorgabe einer digitalisierten Sprachaufnahme und der dazugehörigen Verschriftlichung – die Zeitpunkte der gegebenen Wortformen und deren tatsächliche Aussprache zu bestimmen. Ein solches Programm muss Phoneme aus Kurzzeitspektren, Wörter aus Phonemen und Texte aus Wörtern zusammensetzen können. Darüber hinaus muss es für jedes Wort eine Reihe von Varianten kennen, von denen es jeweils eine aufgrund des vorliegenden Audiosignals auswählen muss.

Um diese Aufgaben bewerkstelligen zu können, umfasst das System u.a. folgende Komponenten (vgl. Abb. 2):

- ein Inventar von Phonemen bzw. Phonen mit dazugehörigen Modellen,
- ein Aussprachewörterbuch⁶ und/oder ein Phonetisierungsprogramm,
- ein Modell für die Darstellung des Textes in Form einer Grammatik bzw. eines Netzwerks,
- ein Mustererkennungsverfahren in Kombination mit einem Signalverarbeitungsprogramm, das die Information der anderen Komponenten berücksichtigt.

Phone unterscheiden sich durch ihre Dauer und ihr Frequenzmuster voneinander. Da jedoch jede Realisierung eines Phons etwas anders ausfällt, scheidet eine deterministische Modellierung aus. In der Praxis ha-

⁶ Unter *Aussprachewörterbuch* wird hier eine zweisepaltige Liste von Einheiten verstanden: die linke Spalte stellt die orthographische Form eines Wortes und die rechte Spalte die entsprechende phonetische Form in SAMPA-Notation dar.

ben sich für die Modellierung solcher Ereignisse spezielle stochastische Automaten (sog. Hidden Markov Modelle) bewährt. Mit Hilfe von Übergangs- und Emissionswahrscheinlichkeiten lassen sich Zeitdauer und Frequenzverlauf modellieren. Vor dem Synchronisations- oder Erkennungsprozess müssen diese HMM trainiert worden sein. Dies geschieht in der Regel nur einmal, nämlich bei der Erstellung eines Aligners oder Erkenners.

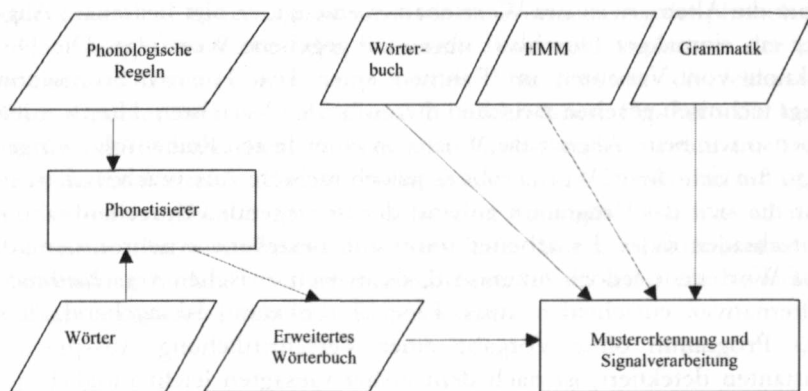


Abb. 2: Architektur des Variantendetektors

Ein Aussprachewörterbuch (in oben festgelegter Bedeutung) zeigt, wie sich Wörter aus Phonemen zusammensetzen lassen. Es hat meist den Vorteil, die phonetische Transkription der Wörter exakter darstellen zu können, als dies ein Computerprogramm kann. Die Nachteile sind die zeit- und kostenintensive Erstellung und der begrenzte Wortschatz eines Wörterbuchs. Insbesondere wenn man alle nur erdenklichen Aussprachevarianten mitberücksichtigen möchte, stößt man schnell an die Grenzen der Realisierbarkeit eines solchen Vorhabens. Es liegt daher nahe, nur solche Wortformen in ein Wörterbuch aufzunehmen, deren Aussprache unregelmäßig ist. Das am IDS entwickelte System SPRAT verfügt sowohl über ein Aussprachewörterbuch als auch über ein Phonetisierungsprogramm (vgl. Schmidt 1996). Letzteres spielt insbesondere bei der Variantendetektion eine wichtige Rolle. Es gestattet mittels kontextsensitiv formulierter Regeln die Generierung von phonetischer Transkription in SAMPA-Notation.

Wie aus der Zusammensetzung von Phonemen ein Wort konstruiert werden kann, so kann aus der Aneinanderreihung von Wörtern ein Text entstehen, wenn dafür bestimmte Regeln erfüllt werden. Man spricht in diesem Zusammenhang auch von einer Grammatik oder einem Sprachmodell (Language Model). Aus der Grammatik wird für die nachfolgende Mustererkennungskomponente ein Netzwerk erzeugt, das den Rahmen für den Erkennungsprozess bildet.

Während ein Spracherkenner den Text durch eine endliche Schleife über die Alternativen des Wortschatzes erzeugt, erfolgt bei einem Aligner ein einmaliger Durchlauf über eine gegebene Wortfolge. Die Detektion von Varianten im Rahmen einer Text-Ton-Synchronisation liegt technisch gesehen zwischen diesen beiden Verfahren. Hier werden ebenso wie beim Aligner die Wörter in einer festen Reihenfolge vorgelegt, für viele dieser Wörter gibt es jedoch mehrere Aussprachevarianten, für die sich das Programm anhand der vorliegenden Sprachaufnahme entscheiden muss. Es arbeitet somit auf Textebene synchronisierend, auf Wortebene jedoch erkennend, da es sich zwischen verschiedenen Alternativen entscheiden muss. Eine Modifikation dahingehend, dass das Programm ohne Vorgabe einer Verschriftlichung Aussprachevarianten detektiert, ist nach dem bisher Gesagten leicht möglich, jedoch ist aufgrund der zahlreicheren Alternativen mit weniger guten Ergebnissen zu rechnen.

Auf die Mustererkennungs- und Signalverarbeitungskomponente soll hier nicht näher eingegangen werden. In SPRAT werden dazu Komponenten von HTK (vgl. Young et. al. 1997) benutzt.

4.2. Projektspezifische Weiterentwicklung

Die Weiterentwicklung des Variantendetektors besteht darin, ihn in die Lage zu versetzen auch feinere Unterschiede bei der Aussprache aufspüren zu können. Dazu müssen

- die Regeln des Phonetisierers erweitert werden bzw. – insbesondere bei Unregelmäßigkeiten – die Varianten in das Aussprachewörterbuch eingetragen werden und
- das Phoninventar um eine Reihe wichtiger Phone erweitert werden

Die Weiterentwicklung eines Aligners zu einem Variantendetektor umfasst daher zwei Phasen:

1. Modellierungsphase,
2. Trainingsphase.

Die 1. Phase dient der phonologischen Modellierung. Zur Zeit verfügt SPRAT über eine Reihe von phonologischen Regeln, die Phänomene im Deutschen wie Elision, Spirantisierung, Fortis-Lenis u.a. modellieren. Dieses Konzept war ursprünglich dafür gedacht, die Text-Ton-Synchronisation möglichst robust hinsichtlich Aussprachevarianten zu machen (vgl. Schmidt/Neumann 1999). Mittlerweile profitiert der Variantendetektor davon.

Die 2. Phase muss durchlaufen werden, wenn sich herausstellt, dass das vorhandene Phoninventar nicht für die Formulierung der phonologischen Regeln ausreicht oder bestimmte Varianten nicht mit der gewünschten Genauigkeit erkannt werden. Dabei werden die fehlenden Phone trainiert bzw. schon vorhandene nachtrainiert. Das setzt die Markierung der entsprechenden Phone in hinreichender Anzahl im Trainingsmaterial voraus. Hierzu erfolgt zunächst ein Detektionslauf über das zu untersuchende Material, dann eine manuelle Überprüfung und Korrektur. Das so aufbereitete Material kann dann als Trainingsmaterial Anwendung finden. In einem zweiten Lauf sollten dann alle Varianten mit der erwarteten Exaktheit erhalten werden.

5. Schluss: COSMAS II als phonetisches Informationssystem

COSMAS II (Corpus Storage and Maintenance System) (vgl. Bodmer 1996, Bodmer et. al. 2002) ist ein Volltextdatenbanksystem, das auch erlaubt, Korpora gesprochener Sprache zu verwalten. Während Transkripte mit einfacher Diskursstruktur (z.B. Interviews, die keinerlei Überlappungen aufweisen) praktisch wie geschriebene Sprache behandelt werden können, bedürfen Transkripte, die mit Kommentaren versehen sind oder Annotationen in der Transkription enthalten (z.B. *LACHT KURZ*) einer zusätzlichen Gliederung. Diese Gliederung – auch Diskursmodell genannt – legt die im Hinblick für diskursanalytische Untersuchungen wichtige Transkriptionsrichtlinie fest, so wie Markierungen wie z.B. Äußerungsgrenzen, Simultanpassagen u.a. Nachdem die Diskurstexte in einem sogenannten Indexierungsprozess in komprimierter Form in die Datenbank aufgenommen und auf einen schnellen Recherchezugriff vorbereitet worden sind, werden in einem separaten Verzeichnis die Audiodaten, die mit den Transkripten aligniert (synchronisiert) worden sind, aufgebaut. Vor dem eigentlichen Recherchieren muss eine Transkriptausswahl getroffen werden. Dazu kann zwischen

einem vordefinierten und einem selbst zusammengestellten Korpus gewählt werden. Als Auswahlkriterien stehen z.Z. zur Verfügung: Aufnahmedatum, Textsorte (z.B. Interview, Schlichtung, Talkshow u.a.), die Korpusbezeichnung und Textfelder.

Für die Datenbank VariaPhonis werden in COSMAS II Kriterien wie Regionalität, Kommunikationssituation u.a. hinzukommen. Eine Reihe der bisher vorhandenen Suchoperatoren steht auch für die Recherche mit VariaPhonis vorgefertigt zur Auswahl. Diese Operatoren können einzeln oder in Kombination verwendet werden. So lassen sich z.B. bereits Wortformen, Prosodie, Diskursgliederung und Sprecherdaten zu Suchanfragen kombinieren. Für VariaPhonis wird künftig noch die vom dem Variantendetektor gelieferte tatsächliche Aussprache in Lautschrift hinzukommen. Letztere wird im Zusammenhang mit der Korpusgliederung nach Regionalität, Textsorte, Kommunikationssituation u.a. die in Abschnitt 3 geschilderten soziolinguistischen Untersuchungen von Variation im empirischen Material gestatten.

6. Literatur

- Bausch, K.-H./N. Berend/W. Kallmeyer/R. Schmidt (2001): Projekt *Aussprachevarianten des Deutschen – Standardsprache und regionale Umgangssprachen*. (Projektbeschreibung: IDS Mannheim, 16.3.2001, 20 S.).
- Bodmer, F. (1996): *Aspekte der Abfragekomponente von COSMAS-II*. In: LDV-Info 8. Informationsschrift der Arbeitsstelle Linguistische Datenverarbeitung. Mannheim: Institut für Deutsche Sprache, S. 112-122.
- Bodmer, F./M. Fach/R. Schmidt/W. Schütte/(2002): *Von der Tonbandaufnahme zur integrierten Text-Ton-Datenbank. Instrumente für die Arbeit mit Gesprächskorpora*. In: Pusch, C.D./W. Raible (Hrsg.): *Romanistische Korpuslinguistik. Korpora und gesprochene Sprache/Romance Corpus Linguistics. Corpora and Spoken Language*. Tübingen: Narr (= ScriptOralia), S. 209-243.
- Duden (2000): *Wörterbuch der deutschen Standardaussprache*. 4., neu bearbeitete und aktualisierte Auflage. Bearbeitet von Max Mangold in Zusammenarbeit mit der Dudenredaktion. Band 6. Mannheim: Dudenverlag.

- Eisenberg, P. (1998): *Das Wort. Der Laut und die Lautstruktur des Wortes*. In: Grammatik der deutschen Gegenwartssprache. 6., neu bearbeitete Auflage. Duden Band 4, S. 17-53.
- Kallmeyer, W. (1997): *Vom Nutzen des technologischen Wandels in der Sprachwissenschaft: Gesprächsanalytische und automatische Sprachverarbeitung*. In: LiLi – Zeitschrift für Literaturwissenschaft und Linguistik 107, S. 124-151.
- König, W. (1998): *Atlas zur Aussprache des Schriftdeutschen in der Bundesrepublik Deutschland*. Bd. 1-2, München: Hueber.
- Kohler, K. (1995): *Einführung in die Phonetik des Deutschen*. Berlin: Schmidt.
- Schmidt, R. (1996): *Konzept und Realisierung eines Phonetisierungsverfahrens*. In: LDV-Info 8. Informationsschrift der Arbeitsstelle Linguistische Datenverarbeitung. Mannheim: Institut für Deutsche Sprache, S. 123-133.
- Schmidt, R./R. Neumann (1999): *Automatic Text-to-Speech-Alignment: Aspects of Robustification*. In: Matousek, V./P. Mautner/J. Ocelíková/P. Sojka (Hrsg.): Text, Speech and Dialogue (Lecture Notes in Artificial Intelligence). Berlin/Heidelberg/New York: Springer, S. 72-76.
- Schmidt, R. (2001): *Instrumente zur Erstellung multimedialer Gesprächskorpora*. In: Lobin, H. (Hrsg.): Sprach- und Texttechnologien in digitalen Medien. Proceedings der GLDV-Frühjahrstagung 2001. Norderstedt, S. 115-127.
- Young, S./J. Odell/D. Ollason/V. Vatchev/P. Woodland (1997): *The HTK Book*. Cambridge University, Entropic Cambridge Research Laboratory Ltd.